

Package ‘simco’

January 2, 2012

Type Package

Title A package to import Structure files and deduce similarity coefficients from them

Version 1.01

Date 2007-04-04

Author Owen Jones <owen.jones@imperial.ac.uk>

Maintainer Owen Jones <owen.jones@imperial.ac.uk>

Depends R(>= 1.8.0), gtools

Description Simco is a package that (1) imports Structure output files into R and (2) carries out similarity coefficient calculations on them.

License GNU General Public License

Repository CRAN

Date/Publication 2007-12-14 16:42:43

R topics documented:

| | |
|----------------------------|---|
| simco-package | 2 |
| SimCoef | 3 |
| SimCoImport | 4 |
| structureexample | 5 |

| | |
|--------------|----------|
| Index | 7 |
|--------------|----------|

simco-package

Simco: a package to carry out similarity coefficient calculations from Structure output files

Description

Simco is a package that (1) imports Structure output files into R and (2) carries out similarity coefficient calculations on them

Details

Package: simco
Type: Package
Version: 1.01
Date: 2007-12-12
License: GNU General Public Licence

There are two functions included in this package:

`SimCoImport`: the import function, designed to import and parse data directly from the Structure output files.

`SimCoef`: the function to do the similarity coefficient analysis.

Examples of Structure output files, and further help, are provided in the docs directory of the simco package (which should be in your R library folder if you have installed simco correctly).

Author(s)

Owen Jones

Maintainer: Owen Jones <owen.jones@imperial.ac.uk>

References

Rebbekc, CA, Jones, OR, & Tsai, IJ (unpublished) SimCo: A program to automate the comparison of multiple structure runs (available from <owen.jones@imperial.ac.uk>)

Pritchard, J. K. et al. (2000) Inference of population structure using multilocus genotype data. *Genetics*. 155, 945-959.

See Also

[SimCoImport](#) [SimCoef](#)

| | |
|---------|---|
| SimCoef | <i>A function to carry out the calculation of similarity coefficients from multiple Structure output files.</i> |
|---------|---|

Description

This function takes multiple Structure files that have already been imported using the "SimCoImport()" function and calculates similarity coefficients for them.

The output is printed to the screen and includes the number of populations (K) that are assumed, the number of individuals (I), the number of Structure runs that have been analysed, the range, median, mean similarity coefficient, and standard error of the mean. Lastly, the output displays the similarity coefficients and the pairs of runs that were compared.

Usage

```
SimCoef(mydata)
```

Arguments

| | |
|--------|---|
| mydata | mydata is a dataframe generated using "SimCoImport()" which is the Structure output files, appended and labelled with a capital letter. |
|--------|---|

Details

An accompanying PDF is given with this package with further details of use.

Value

| | |
|-------------------------------|---|
| Number of populations (K) | The number of populations that were assumed when generating the Structure files |
| Number of individuals (I) | The number of individuals/samples in the datafile |
| Number of Structure runs | The number of Structure runs that have been compared |
| Range | The range of similarity coefficients |
| Median Similarity Coefficient | The median similarity coefficient |
| Mean Similarity Coefficient | The mean similarity coefficient |
| SEM of Similarity Coefficient | The standard error of the mean similarity coefficient |
| The similarity coefficients | a list of the similarity coefficients |

Summary A table showing the pairs of runs that were compared (matches up with the list of similarity coefficients).

Author(s)

Owen R. Jones (<owen.jones@imperial.ac.uk>)

References

Rebbeck, CA, Jones, OR, & Tsai, IJ (unpublished) SimCo: A program to automate the comparison of multiple structure runs (available from <owen.jones@imperial.ac.uk>)

Pritchard, J. K. et al. (2000) Inference of population structure using multilocus genotype data. *Genetics*. 155, 945-959.

See Also

[SimCoImport structureexample](#)

Examples

```
#You should import the files you want to use first using the SimCoImport() command.
#In this example it is assumed that you are using a pre-imported dataset called "structureexample"

data(structureexample)
SimCoef(structureexample)
```

| | |
|-------------|--|
| SimCoImport | <i>A function to import Structure output files from either the current working directory or from a directory of choice</i> |
|-------------|--|

Description

This program imports Structure output files and turns them into a useable dataframe for use in R. Particularly for use by the "SimCoef()" function of this package

Usage

```
SimCoImport(FILE)
```

Arguments

FILE This is a list of the Structure files to be imported.

Details

The list of files is easily generated by putting the files in a directory, using "setwd()" to change the working directory to that directory and then "list.files()" to get the list.

The best way to use this command is by first setting your working directory to one containing the sample files using "setwd()".

Examples of structure output files, and further help, are provided in the docs directory of the simco package (which should be in your R library folder if you have installed simco correctly).

For example:

```
setwd("/Users/orj/Documents/Simco/structurefiles")
myfiles<-list.files()
x<-SimCoImport(myfiles)
x
```

Value

The output is a dataframe with columns of the confidence of assignment to a group and a run identifier (a CAPITAL letter). Rows are individual observations/samples.

Author(s)

Owen Jones

References

Rebbeck, CA, Jones, OR, & Tsai, IJ (unpublished) SimCo: A program to automate the comparison of multiple structure runs (available from <owen.jones@imperial.ac.uk>)

Pritchard, J. K. et al. (2000) Inference of population structure using multilocus genotype data. *Genetics*. 155, 945-959.

See Also

[SimCoef](#)

structureexample *An example dataset for use with SimCo*

Description

This dataset includes 3 runs of Structure (labelled A, B and C) with 3 clusters assumed. Each row represents a tissue sample from a dog. Columns 1 and 2 are simply labels, columns 3-5 are the probabilities that Structure has assigned to membership of the cluster. Column 6 indicates which run the data come from (A, B or C).

Usage

```
data(structureexample)
```

Format

A data frame with 222 observations on the following 6 variables.

Details

A data frame with 222 observations on the following 6 variables. This dataset includes 3 runs of Structure (labelled A, B and C) with 3 clusters assumed. Each row represents a tissue sample from a dog. Columns 1 and 2 are simply labels, columns 3-5 are the probabilities that Structure has assigned to membership of the cluster. Column 6 indicates which run the data come from (A, B or C).

Source

Elaine Ostrander (unpublished data)

References

Pritchard, J. K. et al. (2000) Inference of population structure using multilocus genotype data. *Genetics*. 155, 945-959.

Examples

```
data(structureexample)
```

Index

*Topic **datasets**

structureexample, 5

*Topic **package**

simco-package, 2

SimCoef, 3

SimCoImport, 4

simco (simco-package), 2

simco-package, 2

SimCoef, 2, 3, 5

SimCoImport, 2, 4, 4

structureexample, 4, 5